

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problems Mailbox.**

BEST AVAILABLE COPY

PATENT

IN THE UNITED STATES

PATENT AND TRADEMARK OFFICE

APPLICANTS: Masashi Ito et al.

APPLICATION NO.: 10/052,309

FILING DATE: January 17, 2002

TITLE: APPARATUS AND PROGRAM FOR SEPARATING A
DESIRED SOUND FROM A MIXED INPUT SOUND

EXAMINER: Unassigned

GROUP ART UNIT: 2641

ATTY. DKT. NO.: 20911-06648



CERTIFICATE OF MAILING

I hereby certify that this correspondence is being deposited with the United States Postal Service as first class mail in an envelope addressed to: Box Missing Parts, Commissioner For Patents, Washington, DC 20231, on the date shown below:

Dated: 3/14/02

By: A.C. Smith
Albert C. Smith, Reg. No.: 20,355

BOX MISSING PARTS
COMMISSIONER FOR PATENTS
WASHINGTON, D.C. 20231

TRANSMITTAL OF PRIORITY DOCUMENT

SIR:

Pursuant to the provisions of 35 U.S.C. §119, Applicant is submitting herewith certified copies of the priority applications in Japan, Serial Nos. 2001-016055 and 2001-339622 for filing in the Convention-priority application identified above.

Respectfully submitted,
MASASHI ITO ET AL.

Dated: 3/14/02

By:

A.C. Smith
Albert C. Smith, Reg. No.: 20,355
FENWICK & WEST LLP
Two Palo Alto Square
Palo Alto, CA 94306
Tel.: (650) 858-7296
Fax.: (650) 494-1417

BEST AVAILABLE COPY

#2
11/10/02

日 本 国 特 許 庁
JAPAN PATENT OFFICE

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office

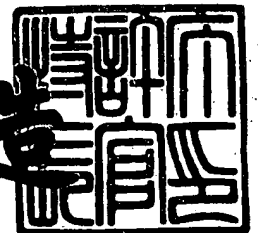
出 願 年 月 日
Date of Application: MAR 21 2002 2001年11月 5日

出 願 番 号
Application Number: 特願2001-339622

出 願 人
Applicant(s): 本田技研工業株式会社

特 許 庁 長 官
Commissioner,
Japan Patent Office

及 川 耕 造



【書類名】 特許願

【整理番号】 H101154601

【提出日】 平成13年11月 5日

【あて先】 特許庁長官 殿

【国際特許分類】 G01H 3/00
G01H 3/08

【発明者】

【住所又は居所】 埼玉県和光市中央1丁目4番1号 株式会社本田技術研
究所内

【氏名】 伊藤 仁

【発明者】

【住所又は居所】 埼玉県和光市中央1丁目4番1号 株式会社本田技術研
究所内

【氏名】 辻野 広司

【特許出願人】

【識別番号】 000005326

【氏名又は名称】 本田技研工業株式会社

【代理人】

【識別番号】 100081721

【弁理士】

【氏名又は名称】 岡田 次生

【選任した代理人】

【識別番号】 100105393

【弁理士】

【氏名又は名称】 伏見 直哉

【選任した代理人】

【識別番号】 100111969

【弁理士】

【氏名又は名称】 平野 ゆかり

【手数料の表示】

【予納台帳番号】 034669

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 混合音の分離装置

【特許請求の範囲】

【請求項 1】 異なる音源より発せられた音響信号と目的信号とが混合された混合入力信号から前記目的信号を分離する音分離装置であって、

前記混合入力信号を周波数分析して各時刻におけるスペクトル及び振幅極大点を計算する周波数分析手段と、

前記スペクトル及び振幅極大点を用いて狭域的な特徴パラメータを分析する狭域層と、前記狭域層により抽出された特徴パラメータを用いて広域的な特徴パラメータを分析する 1 つ以上の広域層とを有し、前記目的信号に関連する特徴パラメータを抽出する特徴抽出手段と、

抽出された特徴パラメータに基づいて前記目的信号を合成する信号合成手段と、を含む音分離装置。

【請求項 2】 前記狭域層及び広域層は、それぞれの層において分析された特徴パラメータを相互に供給し、該供給された特徴パラメータに基づいてそれぞれの層の特徴パラメータを更新する、請求項 1 に記載の音分離装置。

【請求項 3】 前記狭域層は、入力信号に含まれている周波数成分を構成すべき点の候補点の周波数及びその変化と、振幅及びその変化を計算する瞬時符号化層である、請求項 2 に記載の音分離装置。

【請求項 4】 前記広域層は、

前記周波数成分候補点の周波数及びその変化から調波構造を有する周波数成分候補点をグループ化し、該調波構造の基本周波数と該調波構造に含まれる高調波及びその変化を計算する調波性計算層と、

複数の時刻における前記基本周波数及びその変化から信号の連続性を計算するピッチ連続性計算層と、を含む請求項 2 または 3 に記載の音分離装置。

【請求項 5】 前記広域層は前記混合入力信号を発した音源の方向を推定する音源方向推定層をさらに含む請求項 3 に記載の音分離装置。

【請求項 6】 前記広域層は、

前記周波数成分候補点の周波数及びその変化と前記推定された音源方向とから

調波構造を有する周波数成分候補点をグループ化し、該調波構造の基本周波数と該調波構造に含まれる高調波及びその変化を計算する調波性計算層と、

複数の時刻における前記基本周波数及びその変化から信号の連続性を計算するピッチ連続性計算層と、を含む請求項 5 に記載の音分離装置。

【請求項 7】 前記変化として時間変化率を用いる、請求項 3 乃至 6 の何れか 1 項に記載の音分離装置。

【請求項 8】 前記階層はそれぞれ同様の処理を行って特徴パラメータを計算する 1 つまたは複数の計算素子で構成されており、該計算素子は上位の接続する階層及び下位の接続する階層に含まれる各計算素子と前記計算した特徴パラメータを相互に供給する請求項 1 乃至 7 の何れか 1 項に記載の音分離装置。

【請求項 9】 前記計算素子は、上位の接続する階層に含まれる計算素子より供給された特徴パラメータと前記計算した特徴パラメータとの間の整合の度合を示す第 1 の整合性関数を計算し、下位の接続する階層に含まれる計算素子より供給された特徴パラメータと前記計算した特徴パラメータとの間の整合の度合を示す第 2 の整合性関数を計算し、それぞれの整合性関数の積で表される妥当性指標を最大化するように特徴パラメータを更新する、請求項 8 に記載の音分離装置。

【請求項 10】 前記妥当性指標は前記下位の階層に含まれる計算素子に供給される請求項 9 に記載の音分離装置。

【請求項 11】 前記供給された妥当性指標に基づいて閾値を計算し前記妥当性指標の値が該閾値を下回ると計算素子を消滅させる請求項 10 に記載の音分離装置。

【請求項 12】 前記妥当性指標が所定値より大きい場合には下位の階層に新たな計算素子を生成する請求項 10 に記載の音分離装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、混合入力信号から目的信号を分離する信号分離装置に関し、より詳細には、非定常な雑音や複数の音声重畳している混合入力信号から所望の音声信号を分離する装置に関する。

【 0 0 0 2 】

【従来の技術】

従来、センサ等に入力される混合入力信号から必要な音声信号を分離する手法として、スペクトルサブトラクション法が知られている。この手法では、雑音と目的とする信号とが含まれる混合入力信号のスペクトルから雑音のスペクトルを減算することによって目的信号を分離する。

【 0 0 0 3 】

具体的には、例えば、雑音と目的信号とを含む混合入力信号のスペクトルの中から、目的信号が存在せず雑音のみを含むことが明らかな区間から求められるスペクトルを雑音のスペクトルとみなし、これを目的信号の含まれる区間の混合入力信号のスペクトルから減算することによって目的信号のスペクトルを分離する。別の例としては、混合入力信号を集音するマイクロホンとは別に背景雑音のみを集音するマイクロホンを用意し、後者から得られるスペクトルを前者のスペクトルから減算することで混合入力信号から目的信号を分離する。

【 0 0 0 4 】

これとは別の方法として、音声の母音や楽器音の構造上の特徴である調波構造を利用して目的信号のみを抽出する方法がある。一例として、母音の基本周波数とその高調波に対応するスペクトル通過特性を持つ櫛形フィルタを混合入力信号に適用することで、調波構造を持つ信号だけを抽出することができる。

【 0 0 0 5 】

さらに別の方法として、混合入力信号の周波数スペクトルを時系列順に並べた f - t マップにおいて、周波数方向の振幅極大点を走査抽出し、これを周波数成分を構成すべき点の候補点として抽出し、これら極大点を時間方向に順に結んでいくことによって目的信号の周波数成分を抽出する方法が知られている。例えば、 f - t マップ上のある時刻における振幅極大点と次の時刻における振幅極大点とを比較し、周波数やパワー、音源方向などに連続性が見られる極大点同士を時間方向に順次接続していった目的信号を再現する。

【 0 0 0 6 】

複数の信号分離方法を組み合わせた方法もいくつか知られている。特開平9-25

7559号公報には、局所構造情報を利用して振幅極大点をまとめて抽出する手法が開示されている。この手法では、スペクトルの振幅極大点を周波数成分候補点とし、各周波数成分候補点に対し、近傍に位置する近傍点との関係から、周波数成分を構成しているかどうかを判定し、得られた周波数成分候補点に対して、時間、周波数及びパワー値についての連続性を判断し、連続性を持つ点どうしを接続し、周波数成分を抽出する。

【 0 0 0 7 】

【発明が解決しようとする課題】

しかし上述した音分離方法には、それぞれ以下で述べるような問題がある。

【 0 0 0 8 】

まずスペクトルサブトラクション法では、定常的な雑音しか分離することができず、複数の音声信号が重畳する入力信号から1つの音声信号を分離したり、ドアの開閉音のような突発的な雑音を分離したりすることはできない。

【 0 0 0 9 】

櫛形フィルタを用いる方法は、音声信号が定常的な基本周波数を有する場合に有効である。しかし、一般に音声信号の基本周波数は動的に変化するので、実際問題としてこの方法が有効となる場面は少ない。

【 0 0 1 0 】

振幅極大点で周波数成分を抽出する方法においては、振幅極大点の時間方向での連続性を一意に定めることが困難であるという問題がある。特にS/N比が高い場合には、候補となる極大点が増えるため多義性が高くなってしまう。また目的とする信号の周波数成分の近くに他の信号のエネルギーが存在し、振幅極大点が近接している場合には、それらの信号を区別することができない。また、振幅極大点を求めるのに離散フーリエ変換等の手法を用いた場合、入力信号に含まれる音響成分の基本周波数と離散フーリエ変換の解像度が異なったり、入力信号に含まれる音響成分が変調していたり、異なる音源の振幅極大点が近接していると、正確な振幅極大点の周波数を求めることができないので、実際の周波数成分抽出は困難になる。

【 0 0 1 1 】

従って本発明は、非定常な雑音や重畳した複数の音声信号も分離することができる音分離方法を提供することを目的とする。また、目的信号の基本周波数や振幅が動的に変化する場合でも目的信号を分離でき、さらに目的信号と雑音の周波数成分が近接している場合（すなわち S/N 比が高い場合）においても有効に目的信号を分離できる音分離装置を提供することを目的とする。

【 0 0 1 2 】

【課題を解決するための手段】

本発明の音分離装置は、異なる音源より発せられた音響信号と目的信号とが混合された混合入力信号から前記目的信号を分離する音分離装置であって、前記混合入力信号を周波数分析して各時刻におけるスペクトル及び振幅極大点を計算する周波数分析手段と、前記スペクトル及び振幅極大点を用いて狭域的な特徴パラメータを分析する狭域層と、前記狭域層により抽出された特徴パラメータを用いて広域的な特徴パラメータを分析する 1 つ以上の広域層とを有し、前記目的信号に関連する特徴パラメータを抽出する特徴抽出手段と、抽出された特徴パラメータに基づいて前記目的信号を合成する信号合成手段とを含むよう構成される。

【 0 0 1 3 】

この発明によると、特徴抽出手段において狭域的な特徴パラメータと広域的な特徴パラメータを共に扱うことで、目的信号の分離の精度が特定の特徴パラメータの抽出精度に依存することがなくなる。抽出する特徴パラメータには、入力信号に含まれている周波数成分候補点の周波数/振幅値とそれらの変化、調波性、ピッチ連続性、イントネーションなどの時系列データの他に、オンセット/オフセット、音源方向等も含まれる。また特徴抽出手段に設けられる層の数は抽出される特徴パラメータの種類に応じて可変とすることができる。

【 0 0 1 4 】

本発明の別の形態では、前記狭域層及び広域層は、それぞれの層において分析された特徴パラメータを相互に供給し、該供給された特徴パラメータに基づいてそれぞれの層の特徴パラメータを更新するよう構成される。

【 0 0 1 5 】

この形態によると、特徴抽出手段の各層において分析された特徴パラメータを

相互に供給し合うので、特徴パラメータ間で整合性をとることができ、従って特徴パラメータの抽出精度を上げることができる。

【0016】

本発明のさらに別の形態では、前記狭域層は、前記周波数成分候補点の周波数及びその変化と振幅及びその変化を計算する瞬時符号化層である。

【0017】

この形態によると、瞬時の時間変化情報を利用することで同一音源信号の緩やかな振幅、周波数の変化に追従することができる。

【0018】

本発明のさらに別の形態では、前記広域層は、前記周波数成分候補点の周波数及びその変化から調波構造を有する周波数成分候補点をグループ化し、該調波構造の基本周波数及びその変化を計算する調波性計算層と、複数の時刻における前記基本周波数及びその変化から信号の連続性を計算するピッチ連続性計算層とを含む。

【0019】

計算する変化の例としては、時間変化率が挙げられるが、これ以外にも周波数成分候補点の変化を捉えられるものであれば二次導関数等を用いることもできる。

【0020】

この形態によると、非定常な雑音中にある目的信号もその一貫性を利用して分離可能であり、また大局的な特徴パラメータより基本周波数の緩やかな振幅、周波数の変化に追従することができる。

【0021】

本発明の一形態では、前記階層はそれぞれ同様の処理を行って特徴パラメータを計算する1つまたは複数の計算素子で構成されており、該計算素子は上位の接続する階層及び下位の接続する階層に含まれる各計算素子と前記計算した特徴パラメータを相互に供給するよう構成される。

【0022】

この形態によると、抽出する特徴の独立性が高まり、柔軟な特徴パラメータの

更新が実現される。ここで計算素子とは、特徴パラメータに対応して一対一で生成され、それぞれが同様の処理を行い、他の計算素子と特徴パラメータを相互に供給する機能を持つ情報処理素子であり、物理的な素子を意味しているのではない。

【 0 0 2 3 】

本発明のさらに別の形態では、前記計算素子は、上位の接続する階層に含まれる計算素子より供給された特徴パラメータと前記計算した特徴パラメータとの間の整合の度合を示す第1の整合性関数を計算し、下位の接続する階層に含まれる計算素子より供給された特徴パラメータと前記計算した特徴パラメータとの間の整合の度合を示す第2の整合性関数を計算し、それぞれの整合性関数の積で表される妥当性指標を最大化するように特徴パラメータを更新するよう構成される。

【 0 0 2 4 】

この形態によると、計算素子間で相互に特徴パラメータを参照し、特徴パラメータ間の整合性を高めていくことができる。

【 0 0 2 5 】

本発明のさらに別の形態では、前記妥当性指標は前記下位の階層に含まれる計算素子に供給される。

【 0 0 2 6 】

この形態によると、計算素子に対する上位の階層の拘束力を高めて計算の収束時間を短縮したり、逆に拘束力を弱めて影響を軽微にしたりすることができる。これによって計算回数が少ないうちは多くの特徴パラメータを保持しておき、各階層間で整合が取れてくるのに伴って生き残り条件を厳しくし、特徴パラメータの精度を上げるといった制御を行うことができるようになる。さらに、上位層の妥当性指標が更新されるごとに閾値を計算し、妥当性指標の値が該閾値を下回ると計算素子を消滅させることによって不必要な特徴パラメータを早期に取り除くことができ、また前記妥当性指標が所定値より大きい場合には1つ下位の層に新たな計算素子を生成するなどの柔軟なデータ更新が可能となる。

【 0 0 2 7 】

【発明の実施の形態】

以下、本発明の実施形態について図面を参照して説明する。

【0028】

<第1の実施形態>

図1は本発明の第1の実施形態である音分離装置100の全体的な構成を示すブロック図である。音分離装置100には、信号入力部101、周波数分析部102、特徴抽出部103、及び信号合成部104が含まれる。音分離装置100は、雑音や種々音源より発せられる信号が重畳している混合入力信号に含まれる種々の特徴を分析し、特徴間の整合性を整理し、目的信号を分離する。音分離装置100の要部は、例として本発明の特徴を含むソフトウェアを入出力装置、CPU、メモリ、外部記憶装置等を備えたコンピュータやワークステーション等で実行することにより実現されるが、一部をハードウェアにより実現することもできる。図1は、これを踏まえて構成を機能ブロックで表現している。

【0029】

信号入力部101には音分離の対象となる混合入力信号が入力される。信号入力部101は、具体的には例えばマイクロホンのような音響入力端子であり、混合入力信号を直接集音する。この場合音響入力端子は1つとは限らず、2つ以上使用することもできる。音響入力端子が2つ以上ある場合には、後述するように音源方向を目的信号の特徴として利用する形態が実施可能となる。別の実施形態では、混合入力信号は予め準備された音響信号ファイルであり、この場合信号入力部101は音響信号ファイルを取りこむ処理を行う。

【0030】

周波数分析部102は、信号入力部101に入力された信号にA/D変換を施し、デジタル化された信号を適当な時間間隔で周波数分析して各時刻における周波数スペクトルを求め、そのスペクトルを時系列順に並べたf-tマップを作成する。周波数分析は、例えば既知のフーリエ変換やウェーブレット変換、フィルタバンクによる帯域分割などの手法を用いて行う。さらに、求めたスペクトルの振幅極大点を求める。

【0031】

特徴抽出部103は周波数分析部102からf-tマップを受け取り、各スペクトルと

その振幅極大点から特徴パラメータを抽出し、それらの中から目的信号の特徴パラメータを推定する。

【 0 0 3 2 】

信号合成部104は、推定された特徴パラメータから目的信号の波形を再構成する。具体的には、各時刻毎に推定された種々の特徴パラメータから正弦波などのテンプレート波形を使用して目的信号の波形を再構成する。

【 0 0 3 3 】

こうして混合音声信号から抽出され再構成された目的信号は、スピーカ（図示せず）に送られて再生され、またはディスプレイ（図示せず）に送られて目的信号のスペクトルが表示される。

【 0 0 3 4 】

＜特徴抽出部の構成＞

混合入力信号には、入力信号を構成する各音源から発せられた信号の様々な特徴パラメータが含まれている。これらの特徴パラメータはいくつかに分類することができる。例えばピッチや変調、イントネーションといった時間周波数領域に広域的に現れるものや、音源位置情報のように狭域的に現れるもの、周波数スペクトルの極大点やその瞬時変化といった瞬時に現れるものがあり、これらは階層的に表すことができる。また同一音源から発せられた信号の特徴パラメータは、相互に関連性を持っているはずである。本発明ではこれに着目し、特徴抽出部を階層構造とし各階層で異なる特徴パラメータを処理する構成としており、階層間で最も整合が取れるように各階層における特徴パラメータを更新していく。

【 0 0 3 5 】

図2は特徴抽出部103を3層の階層構造とした場合の音分離装置100を示している。図示するように、階層には局所的特徴抽出層106、中間的特徴抽出層107、大局的特徴抽出層108が含まれる。なお、階層構造は抽出する特徴パラメータの種類に応じて4層以上設けることができ、また2層にすることもできる。4層以上となる場合は、中間的特徴抽出層が対応する数だけ増えていくことになる。さらに、一部の階層を並列に配置することもできるが、これについては第2の実施形態及び第3の実施形態と関連して後述する。

【 0 0 3 6 】

特徴抽出部103の各階層は、それぞれ異なる特徴パラメータを分析する。局所の特徴抽出層106と中間的特徴抽出層107、中間的特徴抽出層107と大局的特徴抽出層108とは相互に論理的に接続されている。周波数分析部102において作成されたf-tマップは、特徴抽出部103の中の局所的特徴抽出層106に渡される。

【 0 0 3 7 】

各階層は、まず下位に位置する階層から渡される特徴パラメータに基づいて、自層で抽出する特徴パラメータを計算する。計算された特徴パラメータは上位及び下位に位置する階層に渡される。上下の階層とはお互いの特徴パラメータが拘束条件になっており、接続されている上下層の特徴パラメータと自層の特徴パラメータとの間で整合が取れるように特徴パラメータを更新する。

【 0 0 3 8 】

各層における特徴パラメータと上下層の特徴パラメータとの間で最も整合が取れたとき、特徴抽出部103は最適な解が得られたと判断して、目的信号を再構成できる特徴パラメータを分析結果として出力する。

【 0 0 3 9 】

図3は特徴抽出部103における各階層で抽出する特徴パラメータの組み合わせの一例と、各階層における処理の流れを示すブロック図である。この実施例では、局所的特徴抽出層106では瞬時符号化を、中間的特徴抽出層107では調波性計算を、大局的特徴抽出層108ではピッチ連続性計算をそれぞれ行っている。

【 0 0 4 0 】

瞬時符号化層（局所的特徴抽出層）は、f-tマップに基づいて入力信号に含まれている周波数成分候補点の周波数、振幅、及びそれらの時間変化率を計算する。この計算は、例えば特願2001-16055号に開示されている瞬時符号化方法により実現できる。具体的には、入力信号にA/D変換を施し窓関数を乗じた後、離散フーリエ変換を実行し入力信号のスペクトルを計算する。さらに入力信号のパワースペクトルを計算し、その振幅極大点に対応する単位信号を単数または複数生成する。各単位信号は、周波数、振幅、及びそれらの時間変化率をパラメータとして持っている。各単位信号はA/D変換され離散フーリエ変換によりスペクト

ルが計算される。単位信号が複数の場合、それらは加え合わせられる。入力信号のスペクトルと単位信号和のスペクトルの振幅/位相空間での二乗誤差が計算され、誤差が最小なるよう単位信号の数や各単位信号のパラメータを変更することによって、最終的に入力信号に含まれている周波数成分候補点の周波数、振幅とその時間変化率を得ることができる。

【 0 0 4 1 】

瞬時符号化層には調波性計算層で計算された調波構造の特徴パラメータが入力され、自層で求めた瞬時情報の特徴パラメータとの整合性が検証される。

【 0 0 4 2 】

調波性計算層（中間的特徴抽出層）は、瞬時符号化層で計算された周波数とその時間変化率から、各時刻における信号の調波性を計算する。すなわち、ある基本周波数 f_0 の整数倍($n \times f_0$)の周波数、及びある変化率 df_0 の整数倍($n \times df_0$)の変化率を持つ周波数成分候補点群を1つの調波構造音の周波数成分としてグループ化する。調波性計算層の出力は、調波構造音の基本周波数とその変化率である。調波性計算層にはピッチ連続性計算層で計算された各時刻の基本周波数情報が入力され、自層で求めた特徴パラメータとの整合性が検証される。

【 0 0 4 3 】

調波性計算層は、各時刻において調波構造音を選び出すので、楕形フィルタのように予め基本周波数を記憶しておく必要がない。また、基本周波数変動した場合でも、各時刻において調波構造が存在するので、調波構造音を検出することができる。

【 0 0 4 4 】

ピッチ連続性計算層（大局的特徴抽出層）は、調波性計算層で求められた基本周波数とその時間変化率から、時間的に連続なピッチの流れを計算する。例えば、ある時刻のピッチ周波数とその変化率が得られれば、その前後の時刻のピッチの大まかな値は予測できる。この予測されたピッチと実際にその時刻に存在するピッチとの誤差が一定の範囲内にあるものを、ひとかたまりのピッチの流れとしてグループ化する。ピッチ連続性計算層の出力は、ピッチの流れと、その流れを構成する周波数成分候補点の振幅である。

【 0 0 4 5 】

続いて各階層における処理の流れについて説明する。

【 0 0 4 6 】

まず周波数分析部から得られた f - t マップに対して瞬時符号化計算を行い、特徴パラメータとして入力信号に含まれている周波数成分候補点の周波数 f とその時間変化率 df を算出する(301)。周波数 f と時間変化率 df は、調波性計算層に送られる。

【 0 0 4 7 】

調波性計算層は、各時刻において周波数成分候補点に対応する周波数 f の間の関係及び時間変化率 df の間の関係を調べて、倍音関係にある、すなわち調波構造を持つ周波数成分候補点群をグループ化し（以下「調波性グループ」と言う）、特徴パラメータとして各グループの基本周波数 f_0 及びその変化率 df_0 を求める(302)。この段階では、調波性グループは複数存在し得る。

【 0 0 4 8 】

各時刻で計算された調波性グループの基本周波数 f_0 とその変化率 df_0 は、ピッチ連続性計算層に渡される。ピッチ連続性計算層は、ある一定時間に渡って各時刻における基本周波数 f_0 と変化率 df_0 をそれぞれ比較し、これらを滑らかに結ぶことのできるピッチ連続曲線を推定する(303)。特徴パラメータは、ピッチ連続曲線の周波数とその変化率である。ピッチ連続曲線は、1つの目的信号に雑音等が混入されている場合、1つの f - t マップに対して1つだけ計算されるはずであるが、実際の環境では図4を参照して後述するようにピッチ連続曲線が一意に定まることは少ないので、複数のピッチ連続曲線が候補として推定される。また2つ以上の音声信号を含む混合信号を分離する場合は、2つ以上のピッチ連続曲線が推定されることになる。

【 0 0 4 9 】

こうして調波性計算層、ピッチ連続性計算層で特徴パラメータが求められると、各階層で整合性計算が行われる(304)。具体的には、瞬時符号化層は調波性計算層から特徴パラメータを受け取り、自層の特徴パラメータとの整合性を計算する。調波性計算層は瞬時符号化層とピッチ連続性計算層から特徴パラメータを受

け取り、自層の特徴パラメータとの整合性を計算する。ピッチ連続性計算層は調波性計算層から特徴パラメータを受け取り、自層の特徴パラメータとの整合性を計算する。これらの整合性計算は、各階層で同時並列的に進行する。同時に計算を実行することで、各階層の特徴パラメータ間の整合性をとることができる。

【 0 0 5 0 】

各階層は計算した整合性に基づいて自層の特徴パラメータを更新する。更新された特徴パラメータはさらに図中の矢印で示すように上下の階層に渡され、整合性計算が行われていく (305)。

【 0 0 5 1 】

全ての階層間で整合性がとれた時、計算は終了する (306)。続いて各階層は目的信号の特徴パラメータとして調波構造の基本周波数 f_0 と含まれる高調波 nf_0 (n は整数)、その変化率 dnf_0 、振幅 $a(nf_0, t)$ 及び位相 θnf_0 を各時刻毎に出力する (307)。この結果を使用して信号を再構成することにより、目的音声信号が分離される。このように、様々な特徴パラメータ間の整合性に基づいて全体の計算を並列的に行う手法により、複雑な構造を有する調波構造音をロバストに分離することが可能となる。

【 0 0 5 2 】

上述の説明では、簡単のために f - t マップ上で調波構造をグループ化しているが、このグループ化は瞬時符号化層で抽出される特徴の数に応じて 4 次元以上の特徴空間でも行うことができる。例えば各周波数成分候補点の周波数とその変化率に加えて各周波数成分候補点の振幅とその変化率を利用して、周波数成分候補点の周波数及び振幅の変化が連続するようにグループ化を行うことができる。これは同一音源からの信号のピッチが連続しているのと同様に、同一音源からの信号の振幅が連続していることに対応している。他の瞬時符号化特徴についても同様である。

【 0 0 5 3 】

以上に説明した実施形態のように音声信号の局所構造に注目し音分離を行う手法は、上述の特開平 9-257559 号公報のようにこれまでもいくつか提案されている。このような従来手法で問題となるのは、ある振幅極大点が次の時刻のどの振

幅極大点と結び付くべきかが一意には定まらないことである。この点について、図 4 を用いて説明する。

【 0 0 5 4 】

図 4 は混合入力信号の周波数分析によって得られた f - t マップの例である。混合入力信号は 2 つの連続した音声信号を含み、瞬間的に雑音が入っているものとする。図中の黒丸は混合入力信号のスペクトルの振幅極大点を表す。(a) は従来手法によるピッチ連続性の推定結果を示す。この手法では、ある時刻における周波数方向の振幅極大点を次の時刻における振幅極大点と結び付けていくことで音の流れを推定する。しかし、図示するように接続可能な流れは何通りも考えられ、一意には定まらない。特に S/N 比が低い場合には、目的信号の近傍に結び付ける候補点が増えるので問題はさらに困難となる。

【 0 0 5 5 】

それに対し上述の実施形態では、瞬時符号化によって (b) に示すように離散フーリエ変換解像度のずれや入力信号の変調、周波数成分の近接によって、実際の周波数成分とずれている可能性のある振幅極大点ではなく、周波数成分候補点とその変化率が求まるため、 f - t マップ上で矢印で表しているようにその周波数がどの方向に変化するのかが分かる。従って図 (b) 中に実線と点線で示すように音の流れが明確になり、 \times 印が付されている 2 つの矢印のような周波数成分候補点は雑音として分離される。

【 0 0 5 6 】

さらにこの実施形態では、同一音源より発せられる音声信号に含まれる音響的特徴間には関連性があり、またその性質が急激に変化せず一貫性を持つことに着目している。従って、非定常な雑音中にある音声信号も、音声信号の一貫性を利用して分離可能であり、また大局的な特徴パラメータより同一音源信号の緩やかな振幅、周波数の変化に追従することができる。

【 0 0 5 7 】

また、性質の異なる種々の特徴パラメータを同時に抽出し関連付けることで、単体の特徴抽出精度が確保できない入力信号でも相互に不確定なところを補い合い、全体として特徴抽出精度を上げることができる。

【 0 0 5 8 】

<計算素子>

本発明の実施形態では、各階層は1つまたは複数の計算素子で構成される。本明細書において「計算素子」とは、特徴パラメータに対応して一対一で生成され、それぞれが同一の処理を行い、他の計算素子と特徴パラメータを相互に供給する機能を持つ情報処理素子であり、物理的な素子を意味しているのではない。

【 0 0 5 9 】

図5は各階層の計算素子による構成の一例を示した図である。上から順に大局的特徴抽出層、中間的特徴抽出層、局所的特徴抽出層に対応する計算素子の構成が示されている。ここでは図5の括弧内に示したような上述の実施形態の特徴の組み合わせに対して図5の説明を行うが、他の特徴の組み合わせに対しても同様である。501は周波数分析部により供給されるf-tマップの例であり、4つの時刻 t_1 、 t_2 、 t_3 、 t_4 に対してそれぞれ5、3、5、5個の振幅極大点（図中黒点で表す）が検出された場合を表している。

【 0 0 6 0 】

局所的特徴抽出層では、f-tマップ上の振幅極大点に対応した計算素子が生成される。図5中では、計算素子は黒塗りの正方形（例えば503）で示されている。中間的特徴抽出層では、互いに調波関係にある局所的特徴抽出層の計算素子の1グループに対して1つの計算素子が生成される。図5では、時刻 t_1 、 t_3 、 t_4 についてそれぞれ調波構造が認められるので、中間的特徴抽出層に3つの計算素子j-2、j、j+1が生成される。これらは図中に黒塗りの直方体（例えば504）で示されている。時刻 t_2 に関しては、周波数成分候補点間の数が少なく調波構造が認められるに至らなかったため、この時点では計算素子j-1は生成されない。

【 0 0 6 1 】

大局的特徴抽出層では、調波性計算によって計算された基本周波数とその変化率から、時刻 t_1 から t_4 に渡ってピッチ連続性があると思われるグループに対して計算素子が生成される。図5では、計算素子j-2、j、j+1に対してピッチ連続性があると認められたため、計算素子iが生成されている。これは図7中に横方向に長い1つの直方体（505）で示されている。

【 0 0 6 2 】

整合性計算が進行して計算素子 i の妥当性が強くなると、時間 t_2 に相当する中間的特徴抽出層における計算素子の存在の妥当性が強くなるので、計算素子 $j-1$ が生成される。これは図中白抜きの直方体506で示されている。さらに整合性計算を続けることにより計算素子 $j-2$ 、 $j-1$ 、 $j+1$ の妥当性が強くなると、局所的特徴抽出層において白抜きの正方形（例えば502）で示されている個所の計算素子の存在の妥当性が強くなるので、それぞれ対応する計算素子が生成される。

【 0 0 6 3 】

実際の音分離においては、 f - t マップ上には目的信号以外の音声信号や雑音の振幅極大点があり、これらに対しても局所的特徴抽出層に計算素子が生成され、その中で調波関係のあるグループに対しては、中間的特徴抽出層に対応する計算素子が生成される。特に整合性計算の開始当初は複数の調波性グループが認められることが多い。大局的特徴抽出層においても同様である。しかしこのような計算素子は、整合性計算が進行するにつれ妥当性が低いと判定され、消滅してしまう。このようにして、目的信号の特徴パラメータに対応する計算素子が淘汰されていく。

【 0 0 6 4 】

図5に示した計算素子による各階層の構成は例に過ぎず、また整合性計算の進捗につれ計算素子の構成は常に変化することに注意されたい。なぜなら、上述したように、計算開始時は f - t マップ上の全ての振幅極大点に対して計算素子が生成されるが、計算が進むと妥当性が低い計算素子は消滅し、妥当性の高い計算素子のみが生き残り、計算が収束するからである。図5は各時刻において1つの調波構造しか認められなかった場合、あるいは整合性計算が進行し妥当性の低い調波構造に対応する計算素子が消滅した場合に対応していると考えることができる。

【 0 0 6 5 】

図6は計算素子600の構成の一例を示す機能ブロック図である。以下の説明において、当該計算素子が含まれる階層を N 層、1つ下位の階層を $(N-1)$ 層、1つ上位の階層を $(N+1)$ 層と呼ぶことにする。また $(N+1)$ 層の計算素子の番号

を i 、 N 層の計算素子の番号を j 、 $(N-1)$ 層の計算素子の番号を k で表す。

【0066】

下位整合性計算部604は、 $(N-1)$ 層で計算される特徴パラメータ集合 P_{N-1} より、自層で抽出する特徴に適合したものを見つけ、パラメータ P_{Nj} を計算する。続いて、 N 層の特徴パラメータ P_{Nj} との整合性 R_{Nj} を次式のBottom-Up関数(BUF)により計算する。

【0067】

【数1】

$$R_{Nj} = BUF(P_{Nj}) = \frac{1}{1 + (P_{Nj} - P_{N-1})^2} \quad \dots (1)$$

【0068】

上位整合性計算部601は、上位の $(N+1)$ 層の各計算素子で計算される特徴パラメータの集合 $P_{(N+1)i}$ と、 N 層の特徴パラメータ P_{Nj} との整合性 Q_{Nj} を次式のTop-Down関数(TDF)により計算する。

【0069】

【数2】

$$Q_{Nj} = TDF(P_{Nj}) = \frac{1}{1 + S_{(N+1)i} \cdot (P_{Nj} - P_{(N+1)i})^2} \quad \dots (2)$$

ここで、 $S_{(N+1)i}$ は $(N+1)$ 層の妥当性指標である(妥当性指標については後述する)。

【0070】

パラメータの数は各層にふくまれる計算素子の数に対応している。図6の中間的特徴抽出層にある計算素子の場合は、 $(N-1)$ 層から供給されるパラメータの数は k であり、 $(N+1)$ 層から供給されるパラメータの数は1である。

【0071】

整合性計算部601、604でそれぞれ計算された整合性関数 Q_{Nj} 、 R_{Nj} は乗算部602で掛け合わされて、妥当性指標 S_{Nj} が計算される。妥当性指標 S_{Nj} は、 N 層に

における計算素子 j のパラメータ P_{Nj} の確からしさを表すパラメータであり、パラメータ空間では整合性関数 Q_{Nj} と R_{Nj} との重なり合う部分として表現される。

【 0 0 7 2 】

閾値計算部603は、 N 層内のすべての計算素子のために閾値計算関数(T C F)により閾値 S_{th} を算出する。閾値 S_{th} は、上位層の妥当性指標 $S_{(N+1)i}$ を参照しつつ、計算の初期段階では比較的小さな値に設定され、計算が収束していくにつれ大きな値に設定される。閾値計算部603は、計算素子600には含まれない。

【 0 0 7 3 】

閾値比較部605は、閾値 S_{th} と妥当性指標 S_{Nj} とを比較する。妥当性指標 S_{Nj} が閾値 S_{th} を下回っていた場合は、この計算素子の存在の妥当性が低いことを意味するので、計算素子は消滅する。

【 0 0 7 4 】

パラメータ更新部606は、妥当性指標 S_{Nj} が最大になるようにパラメータ P_{Nj} を更新する。更新されたパラメータ P_{Nj} は、次の計算サイクルで $(N+1)$ 層及び $(N-1)$ 層の計算素子に渡される。

【 0 0 7 5 】

特徴抽出部において最上位にある階層においては、計算素子自体の構成は図 6 に示したものと同一であるが、計算素子に入力されるパラメータは図 7 に示すようになる。この場合、上位層からの妥当性指標の代わりに大局的特徴抽出層にある計算素子のうち最も妥当性の高い素子の指標(S_{win})を使用する。また上位層からのパラメータの代わりに、下位層からのパラメータをパラメータ推測関数($P P F$)607により計算した結果($P_{predict}$)を使用して、整合性 Q_{Nj} 及び閾値 S_{th} を算出する。従って、 $T D F$ は次式のようなになる。

【 0 0 7 6 】

【数 3】

$$Q_{Nj} = TDF(P_N) = \frac{1}{1 + S_{win} \cdot (P_N - P_{predict})^2} \quad \dots (3)$$

【 0 0 7 7 】

妥当性指標 S_{Nj} が高い計算素子は、下位の層 $(N-1)$ 層における計算素子の TCF に与える影響が強くなり、それぞれの妥当性指標を大きくする効果がある。逆に妥当性指標 S_{Nj} が低い計算素子は影響力が小さくなり、 S_{Nj} が閾値 S_{th} を下回ると消滅する。閾値 S_{th} は $(N+1)$ 層の妥当性指標が変化するたびに計算され、さらに TCF は固定でなく計算の進捗により変化する。これにより、計算回数が少ないうちは多くの計算素子(すなわち対応する特徴パラメータ)を残しておき、各階層間で整合が取れてくるのに伴って生き残り条件を厳しくすることができるため、閾値を固定する場合に比べ特徴パラメータの精度を上げることができる。

【 0 0 7 8 】

図 8 は、上述した計算素子により構成された $(N-1)$ 層、 N 層、 $(N+1)$ 層を有する特徴抽出部における計算の流れを説明するフローチャートである。

【 0 0 7 9 】

計算が開始されると、まず必要な初期設定が行われる(801)。続いて、 $(N-1)$ 層、 N 層、 $(N+1)$ 層のそれぞれにおいて、接続する層から入力されるパラメータデータに基づいて、各層の計算素子のパラメータ更新値が計算され(803)、各層の計算素子のパラメータが更新される(805)。さらに妥当性指標の計算も行われる(807)。

【 0 0 8 0 】

計算したパラメータに基づき、各層はそれぞれ接続している層との接続関係を更新する(809)。このとき、妥当性指標が閾値を下回った計算素子は消滅する(811)。また、必要となった計算素子が新たに生成される(813)。

【 0 0 8 1 】

全ての計算素子のパラメータ更新値が設定値を下回ると(815)、各層間の整合が取れたものとして計算を終了する。計算素子の中にパラメータ更新値が設定値を上回るものがあれば、再度更新値が計算され(803)、以下同様の計算が繰り返される。

【 0 0 8 2 】

<第 2 の実施形態>

各階層で抽出する特徴パラメータは第 1 の実施形態に関連して述べた組み合わせに限定されず、採用する特徴のタイプに応じて局所的、中間的、大局的の各特徴抽出層に割り当てる構成を取ることができる。他に利用できる特徴としては、オンセット/オフセットやイントネーション等が挙げられる。それぞれ適当な方法により特徴パラメータが抽出され、整合がとれるように階層間で特徴パラメータがやり取りされる点は、上述の第 1 の実施形態と同じである。

【 0 0 8 3 】

本発明の第 2 の実施形態では、図 9 のように、音響入力端子を 2 つ設けることにより音源方向を特徴として利用するように構成することができる。この場合は、図示するように音源方向分析部 911 を別に設け、音源方向情報を特徴抽出部 915 に供給する。音源方向分析の方法は、周知の技術、例えばマイクロホンに到達する音の時間差または音圧差から音源方向を分析する方法や、あるいは、入力信号を周波数分析し、周波数毎の到達時間差及び/または音圧差から音源方向を分析する等の方法で良い。

【 0 0 8 4 】

音源の方向を分析するために、混合入力信号は、複数の音響入力端子(本実施形態ではマイク L 901 とマイク R 903 の 2 つ)により集音される。周波数分析部 905 では、マイク L 901、マイク R 903 により集音された信号がそれぞれ別々に FFT 等の手法で分析され、周波数スペクトルが求められる。

【 0 0 8 5 】

特徴抽出部 915 には、マイクと同数の瞬時符号化層が設けられる。本実施形態では、マイク L とマイク R にそれぞれ対応する瞬時符号化層(L) 917 と瞬時符号化層(R) 919 が設けられ、スペクトルを受け取る。瞬時符号化層 917、919 は、それぞれ受け取った周波数スペクトルを元に、周波数成分候補点の周波数と振幅、その時間変化を計算する。

【 0 0 8 6 】

瞬時符号化層 917、919 ではまた、調波性計算層 923 にて計算される調波性情報を用いて、計算した周波数成分候補点との整合性を検証する。

【 0 0 8 7 】

音源方向分析部911には、マイク L 901とマイク R 903により集音された混合入力信号が入力される。入力された信号は、音源方向分析部911においてFFTと同じ時間窓の幅で切り出され、2つの信号の相互相関が計算され、その極大点が求められる（図10に示す黒点）。

【0088】

特徴抽出部915には、音源方向推定層921が設けられる。音源方向推定層921は、音源方向分析部911より得られた相互相関のピークのうち、時間方向に引いた線との誤差が一定値より小さいものを、音源方向の違いによる時間差と推定する（図10の場合では、 $\tau 1$ 、 $\tau 2$ 、 $\tau 3$ の3つが推定されている）。このようにして推定された音源方向差による各目的信号の到達時間差は、調波性計算層923に渡される。

【0089】

音源方向推定層921はまた、調波性計算層923から得られる調波性情報の時間差を用いて、推定された各到達時間差との整合性を検証する。

【0090】

調波性計算層923は、瞬時符号化層(L)917と瞬時符号化層(R)919から得られたローカルピークを、音源方向推定層921から得られたそれぞれの到達時間差分ずらして足し合わせることによって、調波性を計算する。具体的には、左右のマイク901、903に、 $\tau 1$ 、 $\tau 2$ 、 $\tau 3$ ずつ到達時間のずれた相似波形の信号が入力されることから、瞬時符号化層917、919それぞれからの出力も $\tau 1$ 、 $\tau 2$ 、 $\tau 3$ だけずれた同じ周波数成分候補点を持つことを利用して、同一音源より到達した目的信号の周波数成分を強調する。このように音分離装置900を構成することによって、複数の目的信号を含む混合入力信号の分離精度を向上させることができる。

【0091】

なお、特徴抽出部915のピッチ連続性計算層925と、信号合成部927の動作は、図3に関して既に説明したと同様である。また各階層が計算素子により構成されることも同様であるが、調波性計算層923における計算素子は、複数の層（すなわち、瞬時符号化層と音源方向推定層）から特徴パラメータを受け取って特徴パ

ラメータを計算し、計算した特徴パラメータを複数の層に渡すように構成される。

【 0 0 9 2 】

＜第 3 の実施形態＞

図 1 1 は、本発明の第 3 の実施形態による音分離装置を示す。

【 0 0 9 3 】

混合入力信号は、複数の音響入力端子（本実施形態では 2 本のマイク L1001 とマイク R1003 の 2 つ）に入力される。周波数分析部 1005 では、マイク L1001、マイク R1003 により入力された信号がそれぞれ別々に F F T 等の手法で分析され、周波数スペクトルが求められる。

【 0 0 9 4 】

特徴抽出部 1015 には、マイクと同数の瞬時符号化層が設けられる。本実施形態では、マイク L とマイク R にそれぞれ対応する瞬時符号化層 (L) 1017 と瞬時符号化層 (R) 1019 が設けられ、スペクトルを受け取る。瞬時符号化層 1017、1019 は、それぞれ受け取った周波数スペクトルを元に、周波数成分候補点の周波数と振幅、その時間変化が計算される。

【 0 0 9 5 】

瞬時符号化層 1017、1019 ではまた、調波性計算層 1023 にて計算される調波性情報を用いて、推定された周波数成分候補点との整合性を検証する。

【 0 0 9 6 】

音源方向分析部 1011 は、周波数分析部 1005 により行われた F F T から、各周波数チャンネルでの相互相関を計算し、その極大点を求める（図 1 2 に示す黒点）。また各周波数チャンネルの音圧差も計算する。

【 0 0 9 7 】

特徴抽出部 1015 には、音源方向推定層 1021 が設けられる。音源方向推定層 1021 は、音源方向分析部 1011 から得られた各周波数チャンネルの信号の相互相関とその極大点、及び各チャンネルの音圧差を併せて求めることにより、極大点を大まかに音源別にグループ化する。このようにして推定された音源方向差による各目的信号の到達時間差は、調波性計算層 1023 に渡される。

【 0 0 9 8 】

音源方向推定層1021はまた、調波性計算層1023から得られる調波構造情報を用いて、推定された到達時間差と音源グループとの整合性を検証する。

【 0 0 9 9 】

調波性計算層1023は、瞬時符号化層(L)1017と瞬時符号化層(R)1019から得られた周波数成分候補点を、音源方向推定層1021から得られたそれぞれの到達時間差分ずらして足し合わせ、さらに音源方向推定層1021から得られる同一音源の情報をを用いることによって、調波性を計算する。

【 0 1 0 0 】

なお、特徴抽出部1015のピッチ連続性計算層1025と、信号合成部1027の動作は、図3に関して既に説明したと同様である。また各階層が計算素子により構成されることも同様であるが、調波性計算層1023における計算素子は、複数の層（すなわち、瞬時符号化層と音源方向推定層）から特徴パラメータを受け取って特徴パラメータを計算し、計算した特徴パラメータを複数の層に渡すように構成される。

【 0 1 0 1 】

【実施例】

上記説明した本発明の第1の実施形態による音分離装置100を用いて、目的信号に雑音を混入した入力信号から目的信号を分離した結果を図13から図15に示す。各図において、(a)は目的信号のスペクトル、(b)は雑音を混入した入力信号のスペクトル、(c)は雑音を分離した出力信号のスペクトルをそれぞれ示している。また各図の横軸は時間(msec)であり、縦軸は周波数(Hz)を表す。入力信号には、ATR音声データベースを使用した。

【 0 1 0 2 】

図13は断続的雑音を目的信号に混入した場合を示す。(a)の目的信号は女性発話「ファミリーレストラン」の一部「ファミリーレス」であり、目的信号の200ms毎に白色雑音を15msずつ混入させたものを(b)の入力信号として用いた。(c)の出力信号は、入力信号から抽出した特徴パラメータから波形を合成して作成した。図より明らかなように、白色雑音はほぼ完全に除去されている。

【0103】

図14は雑音を時間連続的に目的信号に混入した場合の結果である。(a)の目的信号は女性発話「いよいよ」の一部であり、目的信号にS/N比20dBの白色雑音を付加したものを(b)の入力信号として用いた。(c)の出力信号は、入力信号から抽出した特徴パラメータから波形を合成して作成した。目的信号のスペクトル形状が高精度で再現されていることが分かる。

【0104】

図15は他の音声信号を目的信号に混入した場合の分離結果を示す。(a)の目的信号は女性発話「いよいよ」の一部であり、目的信号にS/N比20dBの男性発話「うやまう」を付加したものを(b)の入力信号とした。(c)の出力信号は、入力信号から抽出した特徴パラメータから波形を合成して作成した。(a)の目的信号と比較するとスペクトルは若干異なっているものの、聴感上は問題のないレベルまで目的信号が再現されている。

【0105】

以上本発明を特定の実施形態に関して説明してきたが、本発明はこれらに限定されるものではなく、種々の変更や代替を行うことができる。例えば、説明した各実施形態で使用した特徴パラメータは例示のためであり、今後の研究で発見される新たな特徴パラメータや特徴パラメータ間の関係も本発明において利用することができる。また周波数成分候補点の変化として時間変化率を用いているが、二次導関数等を用いることもできる。

【0106】

【発明の効果】

本発明によれば、混合入力信号のパラメータの時間変化率などの動的特徴量を抽出・利用することで、非定常な雑音が混在する環境下で目的音をロバストに分離することができる。また、予めテンプレートを用意せずに信号の大局的な特徴と局所的な特徴の評価を相互作用させながら同時並列的に処理することで、周波数や振幅が複雑に変化する目的音を柔軟に分離することができる。

【図面の簡単な説明】

【図1】本発明の第1の実施形態による音分離装置の全体の構成を示すブロック

図である。

【図 2】特徴抽出部の階層構造を示す図である。

【図 3】特徴抽出部の各階層で行われる処理を示す図である。

【図 4】従来の方法と本発明の音分離装置によるピッチ連続性の検出を説明する図である。

【図 5】計算素子による特徴抽出部の構成の一例を示した図である。

【図 6】計算素子の一実施形態を示す図である。

【図 7】計算素子の一実施形態を示す図である。

【図 8】図 5 に示した特徴抽出部における処理を説明するフローチャートである。

【図 9】本発明の第 2 の実施形態による音分離装置の構成を示すブロック図である。

【図 1 0】音源方向の推定を説明するためのグラフである。

【図 1 1】本発明の第 3 の実施形態による音分離装置の構成を示すブロック図である。

【図 1 2】音源方向の推定を説明するためのグラフである。

【図 1 3】第 1 の実施形態による音分離装置により音声信号分離を行った結果を示すスペクトル図である。

【図 1 4】第 1 の実施形態による音分離装置により音声信号分離を行った結果を示すスペクトル図である。

【図 1 5】第 1 の実施形態による音分離装置により音声信号分離を行った結果を示すスペクトル図である。

【符号の説明】

100、900、1000 音分離装置

101 信号入力部

102、905、1005 周波数分析部

103、915、1015 特徴抽出部

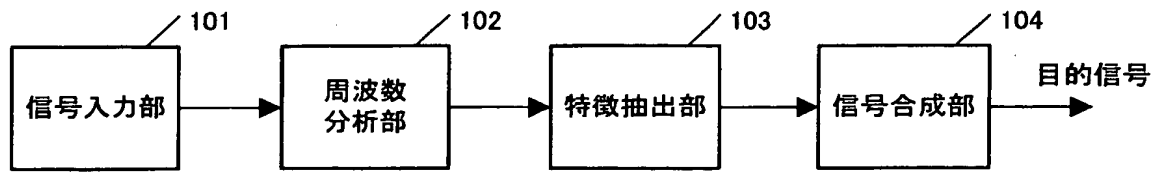
104、927、1027 信号合成部

106、917、919、1017、1019 局所の特徴抽出層（瞬時符号化層）

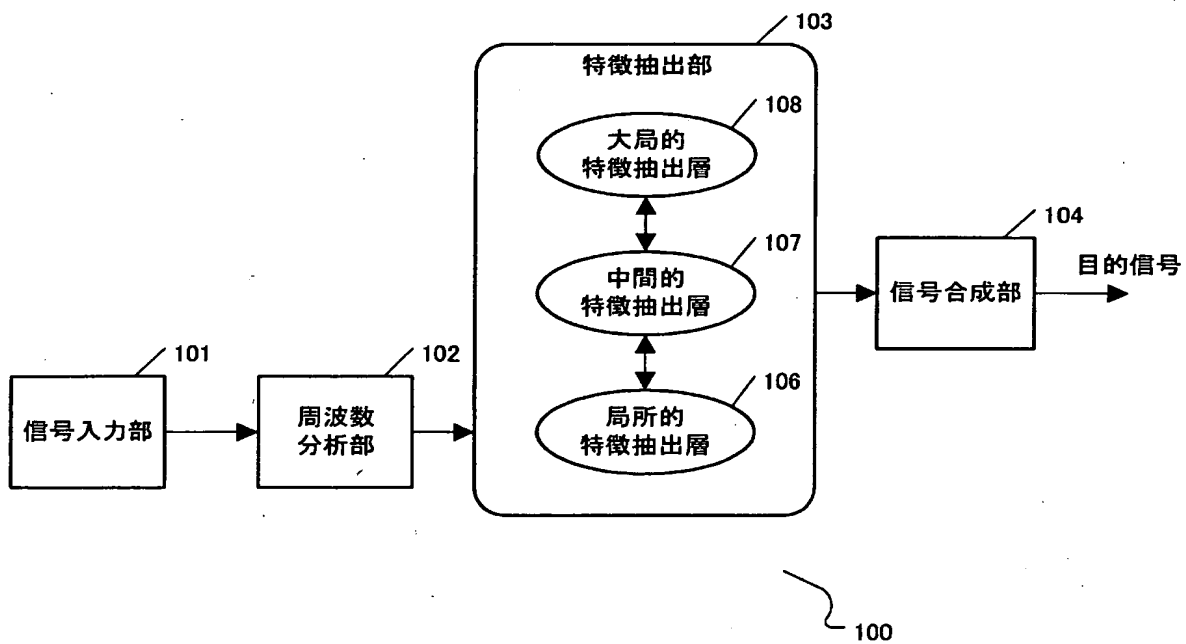
- 107、923、1023 中間的特徴抽出層（調波性計算層）
- 108、925、1025 大局的特徴抽出層（ピッチ連続性計算層）
- 600 計算素子
- 601 上位整合性計算部
- 603 閾値計算部
- 604 下位整合性計算部
- 605 閾値比較部
- 606 パラメータ更新部
- 901、903、1001、1003 マイク
- 911、1011 音源方向分析部
- 921、1021 音源方向推定層

【書類名】 図面

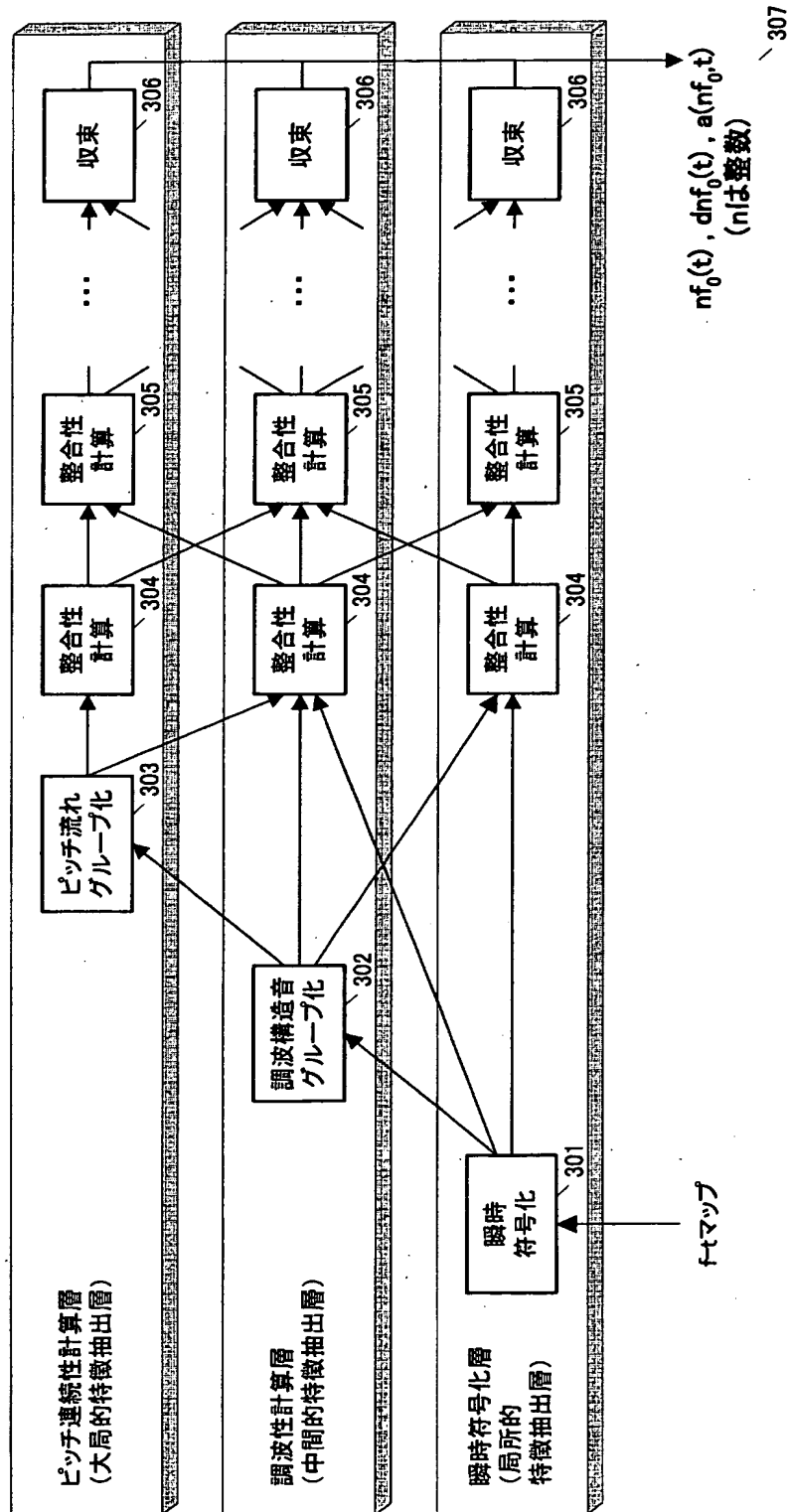
【図 1】



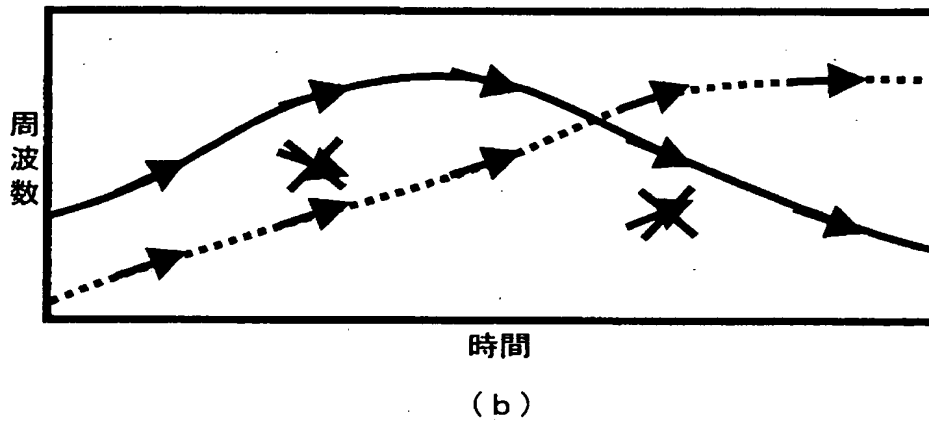
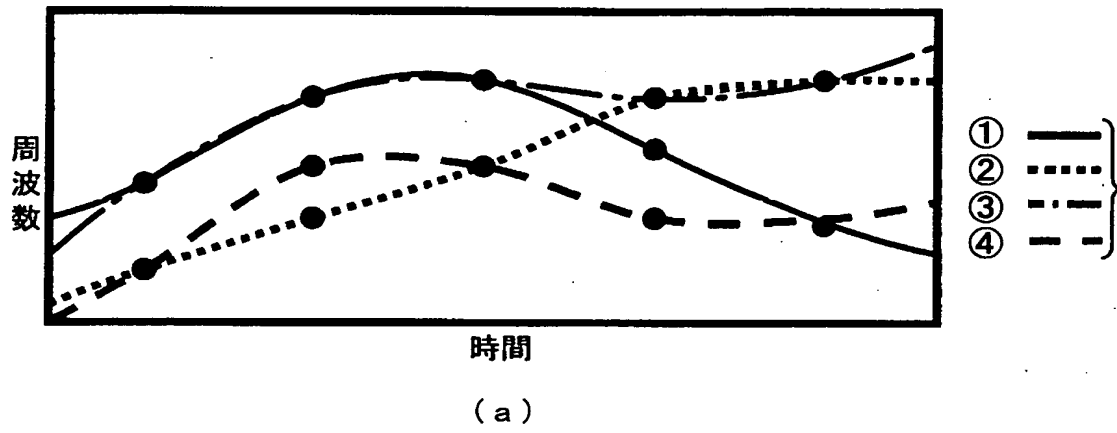
【図 2】



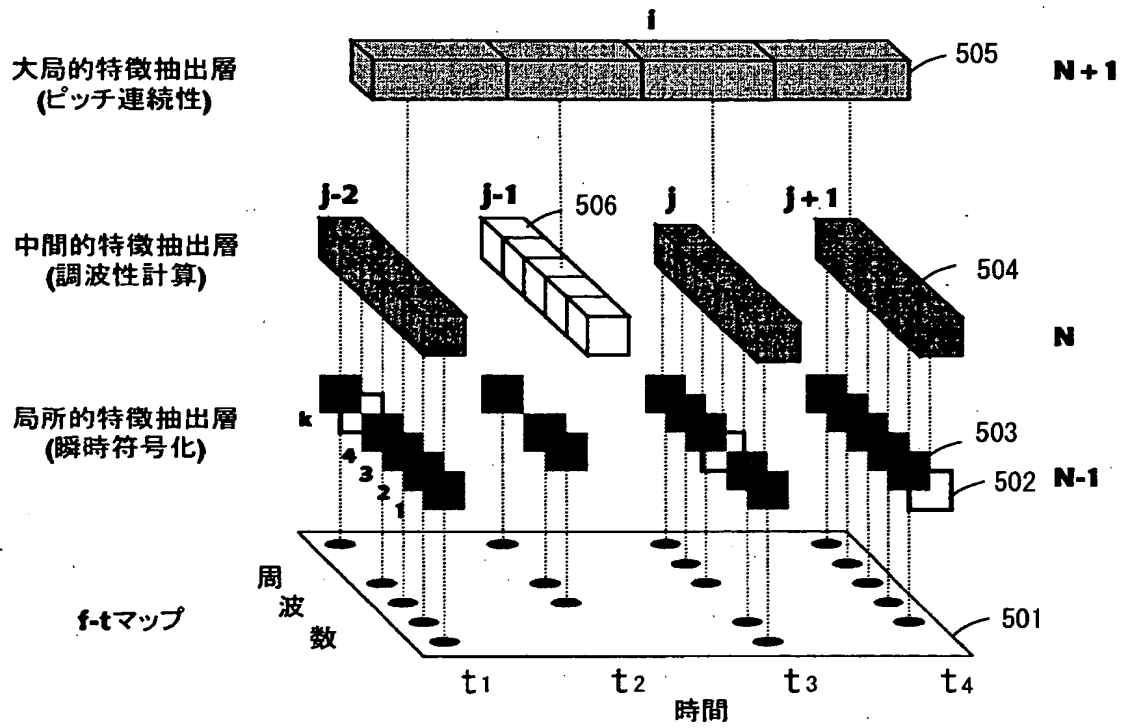
【図3】



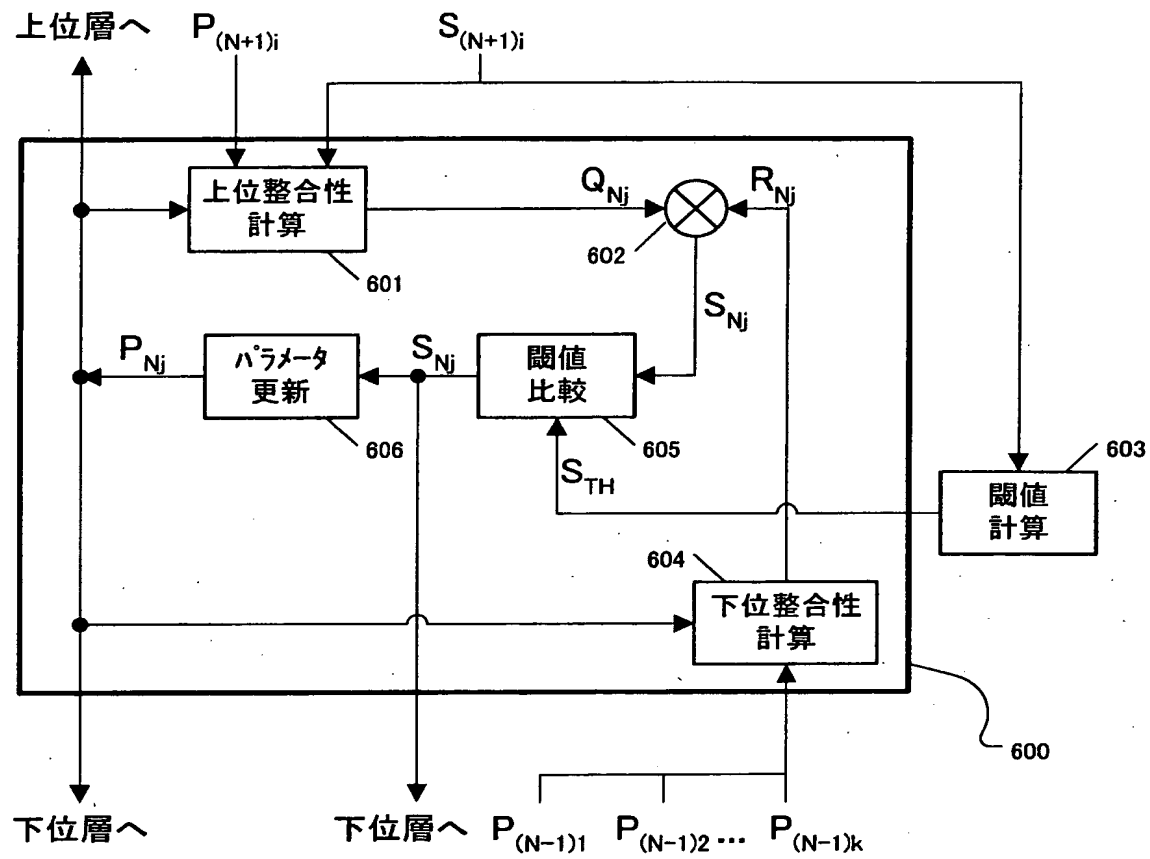
【図 4】



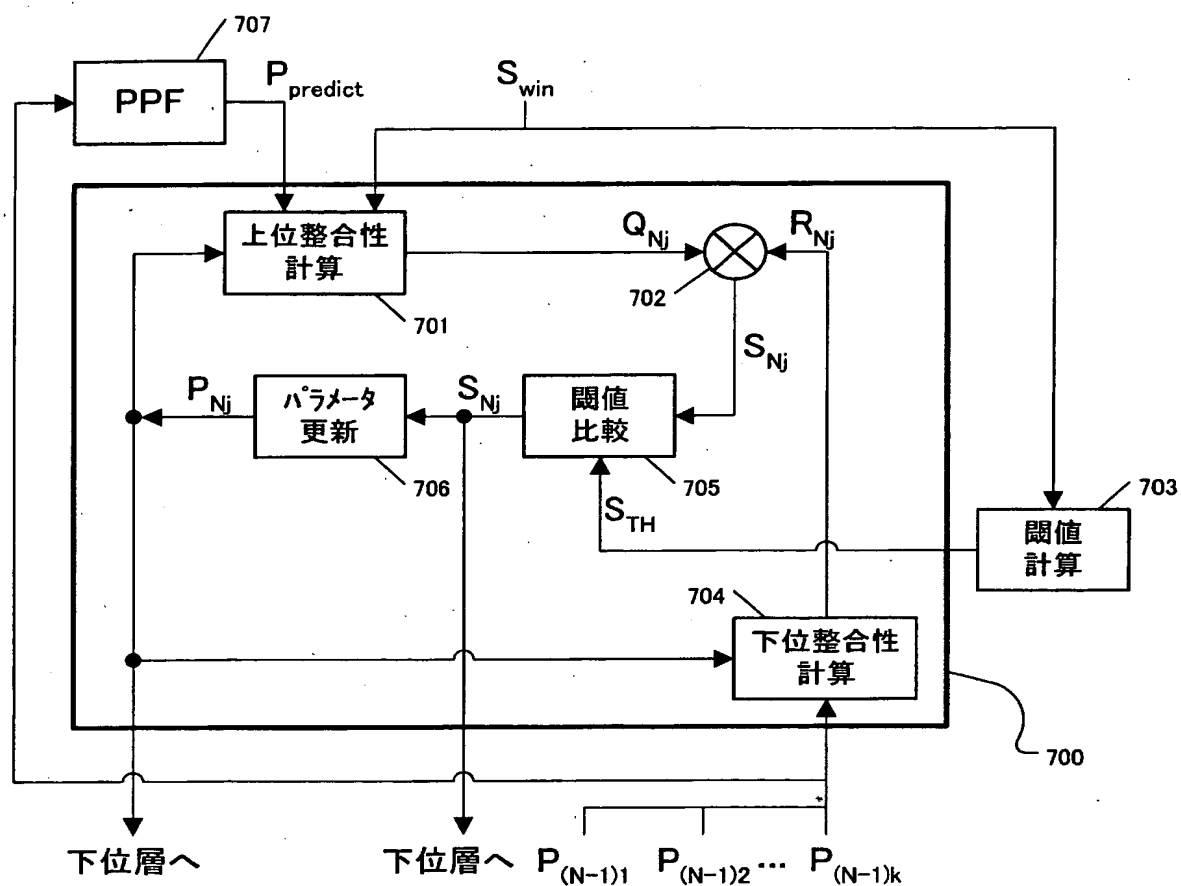
【図 5】



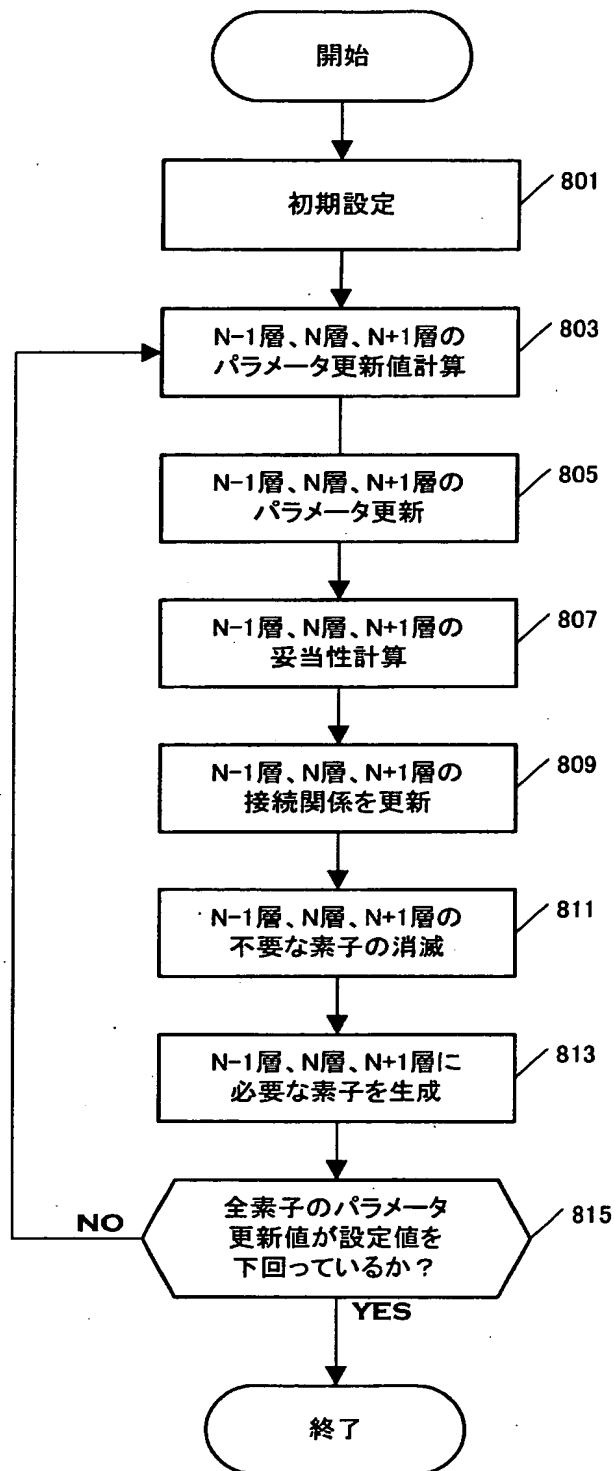
【図6】



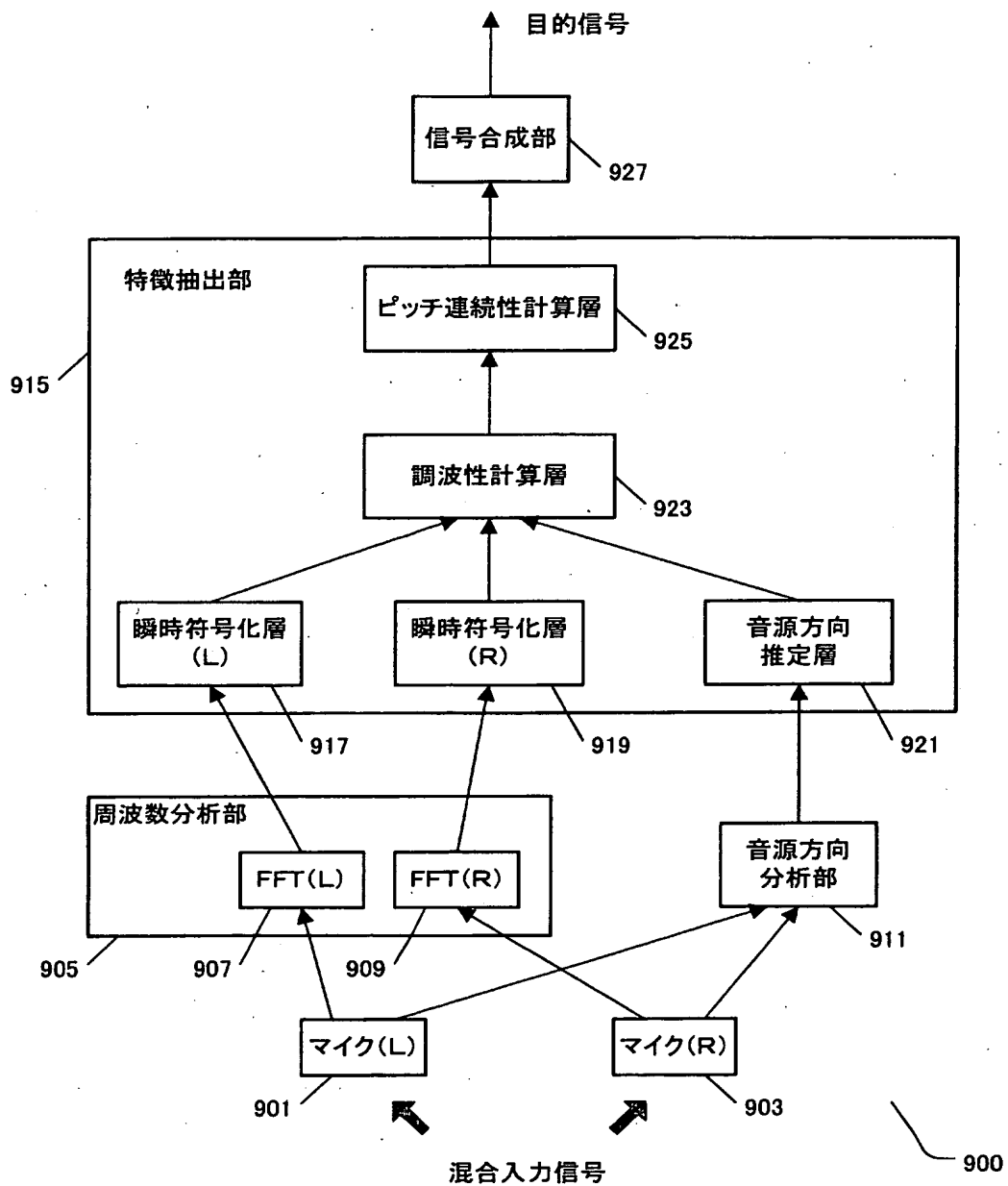
【図 7】



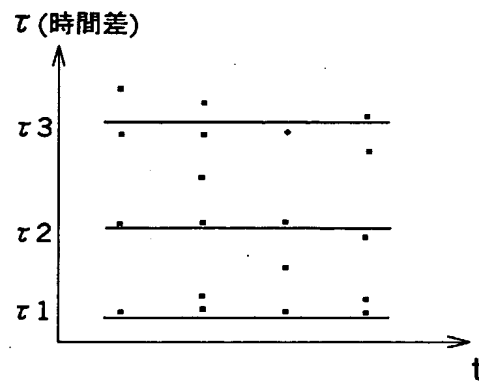
【図 8】



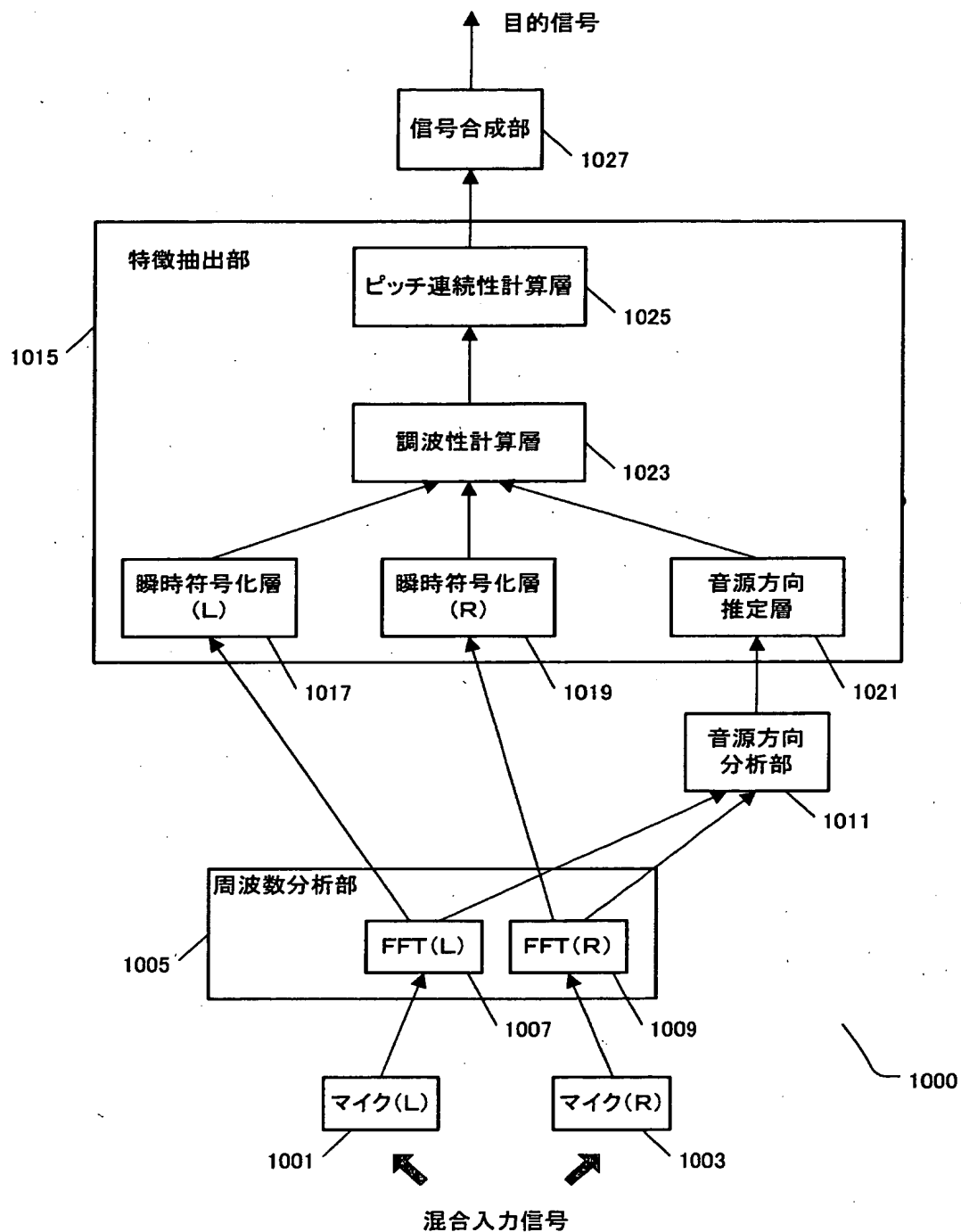
【図9】



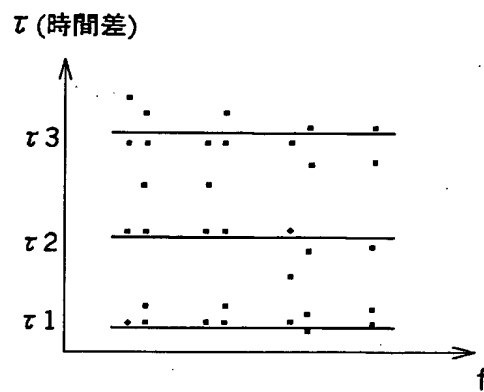
【図 1 0】



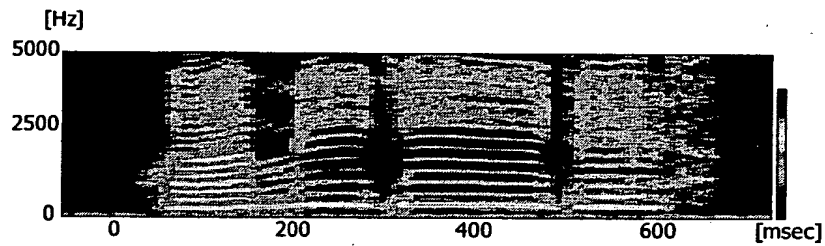
【図 11】



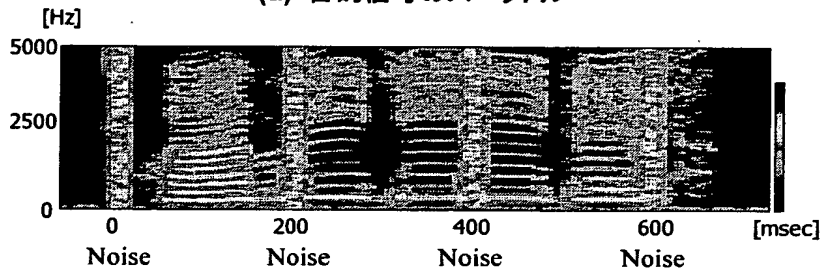
【図 1 2】



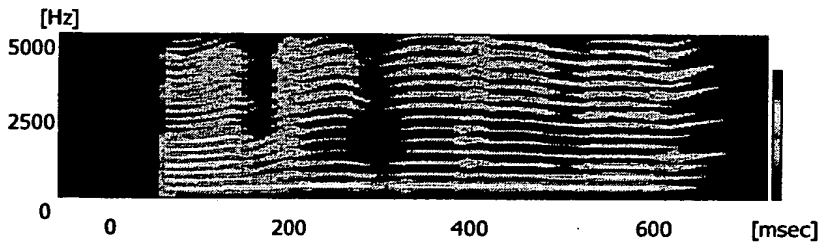
【図 1 3】



(a) 目的信号のスペクトル



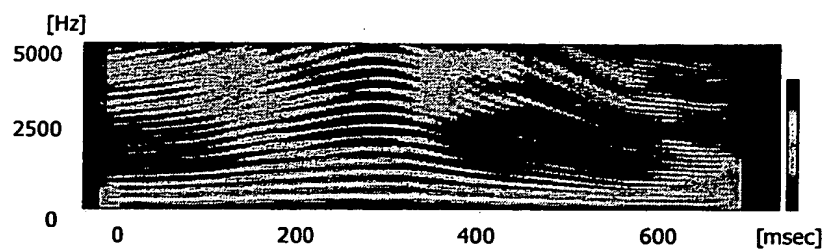
(b) 入力信号のスペクトル



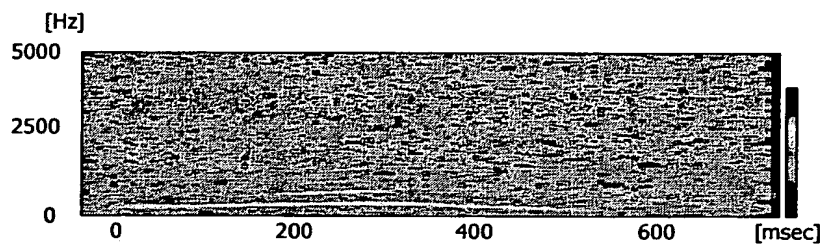
(c) 出力信号のスペクトル

断続的ノイズからの分離

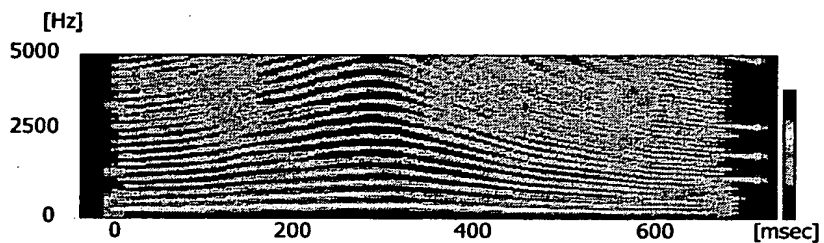
【図 1 4】



(a) 目的信号のスペクトル



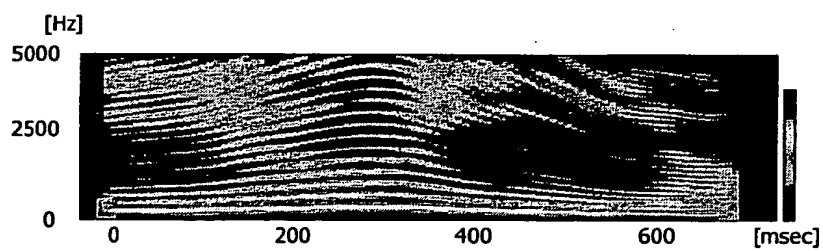
(b) 入力信号のスペクトル



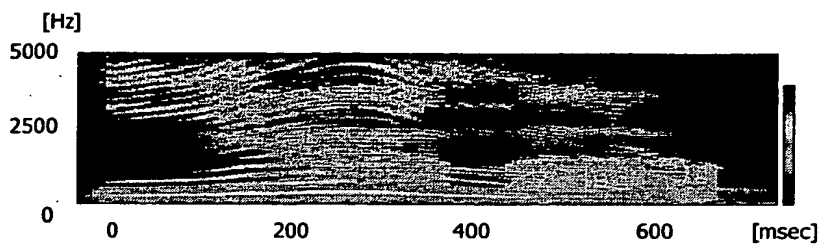
(c) 出力信号のスペクトル

白色ノイズからの分離

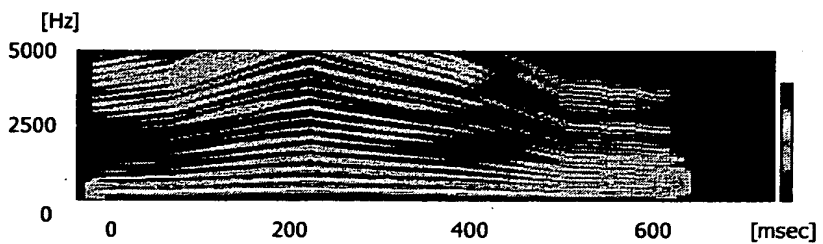
【図 1.5】



(a) 目的信号のスペクトル



(b) 入力信号のスペクトル



(c) 出力信号のスペクトル

他の音声からの分離

【書類名】 要約書

【要約】

【課題】 非定常な雑音や複数の音声重畳している混合入力信号から所望の音声信号を分離する。

【解決手段】 本発明の一実施形態である音分離装置は、異なる音源より発せられた音響信号と目的信号とが混合された混合入力信号を周波数分析して各時刻におけるスペクトル及び振幅極大点を計算する周波数分析手段と、前記スペクトル及び振幅極大点を用いて狭域的な特徴パラメータを分析する狭域層と、前記狭域層により抽出された特徴パラメータを用いて広域的な特徴パラメータを分析する1つ以上の広域層とを有し、前記目的信号に関連する特徴パラメータを抽出する特徴抽出手段と、抽出された特徴パラメータに基づいて前記目的信号を合成する信号合成手段とから構成される。

【選択図】 図 2

出 願 人 履 歴 情 報

識別番号 [000005326]

1. 変更年月日 1990年 9月 6日

[変更理由] 新規登録

住 所 東京都港区南青山二丁目1番1号

氏 名 本田技研工業株式会社